AgDataBox-IA – Web application with artificial intelligence for agricultural data analysis in precision agriculture

Claudio Leones Bazzi[1], Wendel K. M. Oliveira[1], Ricardo Sobjak[1], Kelyn Schenatto[1]

[1]Federal University of Technology – Paraná /UTFPR - Computer Science Department, Av. Brazil, 4232 - 85884-000 - Medianeira, PR – Brazil.

A paper from the Proceedings of the
16th International Conference on Precision Agriculture
21-24 July 2024
Manhattan, Kansas, United States

Abstract.
Agriculture has been continually evolving, incorporating hardware, software, sensors, aerial surveys, soil sampling for chemical, physical, and granulometric analysis (based on sample grids), and microclimatic data, leading to a substantial volume of data. This requires platforms to store, manage, and transform these data into actionable information for decision-making in the field. In this regard, Artificial Intelligence (AI) is the most widely used tool globally to mine and transform vast data volumes into definitive answers for decision-making. To address this, the Precision Agriculture Laboratory, based at the Federal University of Technology – Paraná, Campos Medianeira, has been conducting data storage and mining studies to convert them into crucial information for decision-making in agricultural fields. Therefore, the project will encompass two components: 1) the development of a computational module (AgDataBox-AI) that allows for maintaining a historical organization of crops, field operations, management, harvest, data collected from various sources, among others, and which can standardize the data bank maintained by the application to ensure that such data can be used as a data source for the application of artificial intelligence algorithms, and 2) the development of an AI module capable of assisting farmers and technicians with accurate information for managing agricultural practices such as fertilizer application, irrigation, soil management, cultural practices, and the application of agricultural pesticides. From this perspective, this research aims to optimize the use of agricultural inputs and field activities, focusing on mitigating factors harmful to agricultural productivity in the Western region of Paraná, achieving significant gains in reducing environmental impacts and enhancing the profitability of agricultural activity.

Keywords.
Precision agriculture, Sustainability, soil attributes, IA, AgDataBox-Map

Bazzi, C. B., Oliveira, W. K. O., Schenatto, K., Sobjak. R. (2024). AgDataBox-IA – Web application with artificial intelligence for agricultural data analysis in precision agriculture. In Proceedings of the 16th International Conference on Precision Agriculture (unpaginated, online). Monticello, IL: International Society of Precision Agriculture.

Introduction

Agriculture has been continually advancing, incorporating hardware, software, sensors, aerial surveys, soil sampling for chemical, physical, and granulometric analysis (based on grid sampling), and microclimatic data, leading to vast data. These data demand platforms that can store, manage, and convert them into actionable insights for decision-making in the field. In this scenario, Artificial Intelligence (AI) stands out as the most utilized tool globally to mine and transform large data volumes into tangible responses for decision-making. To this end, the Precision Agriculture Laboratory, housed at the Federal University of Technology – Paraná, has been conducting data storage and mining studies to transform them into crucial insights for decision-making in agricultural fields. Therefore, the aim of this research will be to 1) develop a computational module (AgDataBox-AI) that can maintain a historical organization of crops, field operations, management, harvest, and data collected from various sources, among other things, and standardize the data bank maintained by the application, ensuring these data can serve as a data source for the application of AI algorithms, and 2) develop an AI module capable of assisting farmers and technicians with accurate information for the management of agricultural practices such as fertilizer application, irrigation, soil management, cultural practices, and the application of agricultural pesticides.

Materials and methods

For the construction of ADB-AI, it will be essential to use the AgDataBox API, in which new functionalities will be incorporated, taking into account a historical aspect of soil data, climate, culture, cultivar, vegetation indices, field operations, harvest, and management carried out. Focusing on data management, the aim is to keep them organized and standardized, making them readily available for the application of AI algorithms (Figure 1).

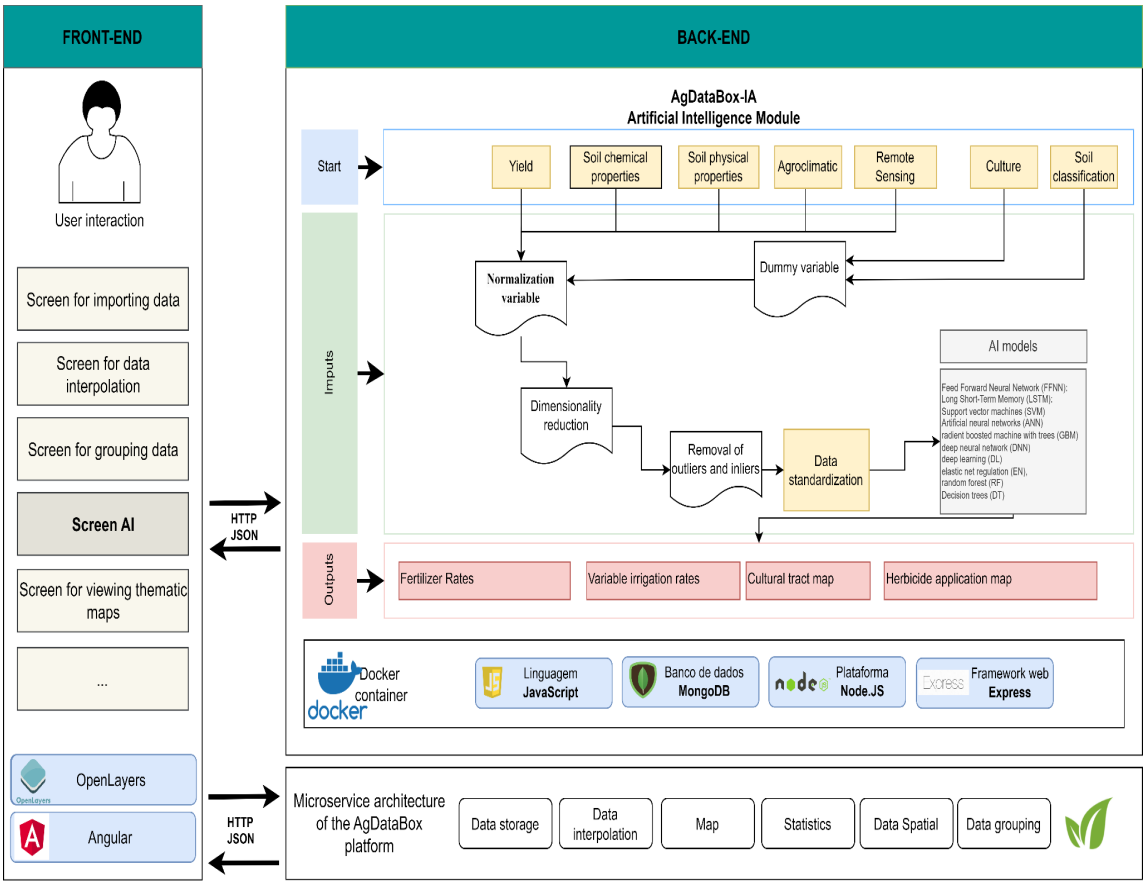

Figura 1 - Fluxograma do módulo de IA.

For data standardization, four methodologies present in the literature and accepted by the scientific community will be used, namely: Normalization: For the application of this technique, the computational module developed by Schenatto et al. (2017); will be used; Outlier treatment: for this step of data standardization, the ADB-Cleaning computational module (Alves et al. 2023); will be employed; Encoding of categorical variables: converting qualitative data into quantitative data, with methods: 1) Label encoding, which is used to convert categorical columns into numeric ones; and 2) one-hot encoding, where the process occurs in two ways, splitting categories into different columns, being 0 for others and one as an indicator for the appropriate column (Dutt et al. 2021). This step will be applied to nominal parameters, such as culture, soil type, climatic classification, and finally, Dimensionality reduction: In some cases, the data may have many features or variables, leading to performance problems or model overfitting. In these situations, dimensionality reduction techniques such as Principal Component Analysis (PCA) or Feature Selection are recommended, which can be applied to reduce the amount of data without significant information loss.

Software Component

In web application development, various tools are used to build complete and efficient systems, both on the front-end and the back-end: i) Front-end: Angular will be used as a framework, and for viewing interactive maps, OpenLayers will be employed as a JavaScript library; and ii) Back-end: Node.js is a JavaScript code execution environment. The Express framework is widely used with Node.js, providing an abstraction layer that facilitates HTTP request routing, session management, cookies, and other standard features. MongoDB will be used for storing user-provided data, which is a NoSQL, document-oriented database that offers schema flexibility and high scalability. MongoDB is widely used with Node.js and Express, allowing easy and rapid integration between the database and the back-end server (Pushkarev and Yakubailik 2021).

Results and Discussion

With the data available in the application, interpolation methods available on the AgDataBox platform should be used to generate continuous surfaces for each attribute. AI algorithms will then be applied for each cultivated area to identify productivity patterns based on soil, plant, and climate attributes and generate similar management zones based on the studied traits. In the event of developing Management Zones, it will be validated through analysis of variance (ANOVA), with the Tukey test applied for multiple comparisons (with a significance level of $p<0.05$). In addition, indices will be used to evaluate the management zones, including the Fragmentation Index and the Variance Reduction Index.

Conclusion or Summary

From this perspective, this research aims to optimize agricultural inputs and field activities, focusing on mitigating factors harmful to agricultural productivity in the Western region of Paraná. The study seeks significant gains in reducing environmental impacts and improving the profitability of farming activities.

Acknowledgments

The authors are grateful to the Federal University of Technology - Paraná (UTFPR), the Araucária Foundation (Fundação Araucária), the Coordination for the Improvement of Higher Education Personnel (CAPES), the National Council for Scientific and Technological Development (CNPq), the Itaipu Technological Park Foundation (FPTI), AgriLab, and The Brazilian Agricultural Research Corporation (Embrapa), for the support received.

References

Alves, F. de L., de Souza, E. G., Sobjak, R., Leones Bazzi, C., Angel Uribe Opazo, M., Marcos Massao Hachisuca, A., & Mercante, E. (2023). Techniques for Cleaning Yield Data for the Creation of Thematic Maps. SSRN Electronic Journal. https://doi.org/10.2139/SSRN.4360372

Dutt, R., Rathi, P., & Krishna, V. (2021). Novel mixed-encoding for forecasting patent grant duration. World Patent Information, 64, 102007. https://doi.org/10.1016/J.WPI.2020.102007

Pushkarev, A. A., & Yakubailik, O. E. (2021). A web application for visualizing, analyzing, and processing agricultural monitoring spatial-temporal data. http://ceur-ws.org. Accessed 4 August 2023

Schenatto, K., de Souza, E. G., Bazzi, C. L., Gavioli, A., Betzek, N. M., & Beneduzzi, H. M. (2017). Normalization of data for delineating management zones. Computers and Electronics in Agriculture, 143, 238–248. https://doi.org/10.1016/j.compag.2017.10.017